Remarks

This amendment is responsive to the Office action mailed October 6, 2003 in connection with the above-identified patent application. In that Action, the finality of the previous Office Action was withdrawn pursuant to 37 CFR § 1.114, and Applicant's submission filed on August 29, 2003 was entered.

Claims 55-59, 61-65, and 67-71 were rejected under 35 U.S.C. § 103(a) as being unpatentable over U.S. Patent No. 5,727,199 to Chen, et al. in view of U.S. Patent No. 5,647,058 to Agrawal, et al.

Claims 60, 66, and 72 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Chen, et al. in view of Agrawal, et al. and further in view of U.S. Patent No. 6,094,651 to Agrawal, et al.

The Present Application:

For purposes of review, the present application relates to a method, an apparatus, and an article of manufacture providing for the creation and use of an index to access a subject multi-dimensional database. An embodiment of the invention receives an indication of a number of features of a subject multi-dimensional database to be identified, and then performs feature identification on the subject multi-dimensional database. An index for accessing the subject multi-dimensional database is created using the identified number of features. The index itself may be a multi-dimensional database.

As described in the specification on page 11, beginning at line 18:

> In order to create the index, the Index System 124 passes the indexing parameters 128 to feature identification software 122 for use in

> finding "features" in the subject multi-dimensional database 136. Then, the Index System finds the "features" with the feature identification software 122. Next, the Index System builds the index 134 using the features as points for indexing into the subject multi-dimensional database 136. Then, the Index System 124 provides navigation capabilities for navigating the index 134 to access the subject multi-dimensional database 136.

The invention provides a user interface to set up definitions for the subject multi-dimensional database to be mined, dimensions to be mined, measures to be mined, mining technique (i.e., feature identification) parameters, and number of results to be stored. The user interface directly drives a mining run. Additionally, the invention supports traversal of the multi-dimensional database, execution of the mining technique, and generation of result data. In one embodiment, the mining technique scans the subject multi-dimensional database and the result data is used to create an index. Operations for creating and deleting the index are provided as well as for outline definition, data population, cell note creation, and link partition definition. The invention provides capabilities for exploration and visualization of the result data against the subject multi-dimensional database.

Overall, the invention automatically builds the index, along with links to the subject multi-dimensional

database. The invention also stores the index data in a spreadsheet data file, so that a spreadsheet user could view a list of deviations in one spreadsheet and link to cells in the subject multi-dimensional database using a linked partition mechanism.

There are many advantages to the invention including its straightforward implementation. Also, the invention does not require any additional functions or support from the developers of the subject multi-dimensional database, it does not modify the existing subject multi-dimensional database, and it does not store extra data in the subject multi-dimensional database. Extra explanations are selectively stored as cell notes on the index as desired, providing visualization and navigation of the multi-dimensional data. Furthermore, the invention is easily managed and can be applied with any data mining technique that can identify point of interest in a multi-dimensional database (i.e., a feature identification technique). The infrastructure of the invention supports plug-in techniques and can extend the solution beyond deviation detection.

U.S. Patent No. 5,727,199 to Chen, et al.:

U.S. Patent No. 5,727,199 to Chen, et al. teaches a system, two-step method and program which, given a large training set of data tuples, in a first phase, called the feature identification phase, identifies features, which have good power in separating data tuples, based on a subset of the training set. In a second phase, called the feature combination phase, the identified features are evaluated in combination against the entire training set to determine final classification rules.

U.S. Patent No. 5,647,058 to Agrawal, et al.:

U.S. Patent No. 5,647,058 to Agrawal, et al. teaches a high dimensional indexing method which takes a set of objects that can be viewed as N-dimensional data vectors and builds an index which treats the objects like k-dimensional points. The method first defines and applies a set of feature extraction functions that admit some similarity measure for each of the stored objects in the database. A feature vector is then transformed in a manner such that the similarity measure is preserved and that the information of the feature vector is concentrated in only a few coefficients. The entries of the feature vectors are truncated such that the entries which contribute little on the average to the information of the transformed vectors are removed. An index based on the truncated feature vectors is subsequently built using a point access method (PAM).

U.S. Patent No. 6,094,651 to Agrawal, et al.:

U.S. Patent No. 6,094,651 to Agrawal, et al. teaches a method for locating data anomalies in a k-dimensional data cube that includes the steps of associating a surprise value with each cell of a data cube, and indicating a data anomaly when the surprise value associated with a cell exceeds a predetermined exception threshold.

All Pending Claims are Patentably Distinct and Unobvious Over the References of Record:

Referring once again to the Office Action in greater detail, claims 55-59, 61-65, and 67-71 were rejected as being unpatentable over U.S. Patent No.

5,727,199 to Chen, et al. (Chen) in view of U.S. Patent No. 5,647,058 to Agrawal et al. (Agrawal[1]). The Examiner took the position that Chen teaches a method, apparatus, and article of manufacture of accessing a subject multi-dimensional database stored on a data store connected to a computer.

Applicants respectfully disagree with the Examiner's understanding of Chen. The Examiner took the position that Chen teaches an apparatus and an article of manufacture "accessing a subject multi-dimensional database stored on a data store connected to a computer, comprising:

a) receiving an indication of a number of features of said subject multi-dimensional database to be identified (col. 3, lines 56-61 and col.4, lines 30-34);

b) performing feature identification to identify the indicated number of features (col. 4, line 60 - col. 7, line 30)..."

However, applicants respectfully submit that the Chen does not teach the above as suggested by the Examiner. Further, applicants submit that the Examiner has not met her prima facie burden of proof of anticipation as required by the law and rules. In support of applicants' position, and for the convenience of the Examiner and for the record, applicants have prepared a table for side-by-side comparison of independent claim 55 against portions of the Chen patent relied upon by the Examiner as a reference anticipating the claim. As can be seen, there is no correspondence between the two. It is respectfully submitted that the claims are not anticipated or rendered obvious by the Chen patent.

Claim 55	Chen, et al. '199

Receiving an indication of a number of features of said subject multi-dimensional database to be identified	"The object of the invention is to identify features from $\{A_1, A_2, \ldots, A_n\}$...", however, there is no limitation with regard to an indication of the number of features to be identified
Performing feature identification to identify the indicated number of features	A method is described for identifying features, e.g. Gender, Age, Beverage and State, in a subset of a training database E

It should be noted that Chen defines features as the n attributes of the database tuple $\{A_1, A_2, \ldots, A_n\}$ (col. 3, lines 48-50). Features, as described in the present application, however, are more complex in nature. For example, an exemplary scenario is described starting on page 13, line 12 where the object is to obtain the three most prominent features. On page 14, lines 9-17, a table is provided illustrating the data returned by the feature identification technique of the particular embodiment being described. The result, as described in lines 1-4 on the same page, is an ordered list of multi-dimensional points, identifying regions of interest.

As can be seen from the above, the Chen patent does not teach, suggest, or fairly disclose receiving an indication of a number of features, and performing feature identification to identify the indicated number of features of the type described in the instant application.

For at least the above reasons, it is respectfully submitted that independent claims 55, 61 and

67, and dependent claims 56-59, 62-65 and 68-71, depending respectively therefrom, are patentably distinct over the references of record.

In addition to the above, the Examiner cites the Agrawal[1] patent as teaching high-dimensional indexing by taking N-dimensional data vectors and building an index of k-dimensional points (col. 4, lines 6-30). This is not relevant to the present application, however, as the application describes an index having the same dimensions as the subject database (page 15, lines 14-16 and lines 25-27). It should be noted also that the Agrawal[1] patent is directed toward a multimedia database storing images, audio and video components (col. 1, lines 58-60). Combining the Agrawal[1] patent with the Chen patent would not provide features as claimed in the present application.

Regarding claims 56, 62 and 68, the Examiner took the position that the k-dimensional index taught by Agrawal[1] (col. 4, lines 6-9) was the same as creating the index of the present application, wherein creating the index comprises creating a multi-dimensional database that is derived from the subject multi-dimensional database. Each of the aforementioned dependent claims includes a limitation such that the index is itself a multi-dimensional database. Applicants show below, however, that merely separating and index from the data in a multi-dimensional database does not make the index itself into a multi-dimensional database, and there is no teaching or suggestion in Agrawal[1] that the index is a multi-dimensional database.

There is a clear lexical distinction between an index and a database in the art. As defined in the *Dictionary of Computing*, Oxford University Press, 1990 (Appendix I), a database is normally and strictly, a data file that is defined and accessed using the facilities of a

database management system (DBMS). This implies in particular (a) that it is defined by means of a schema that is independent of any programs that access the database and that it uses direct access storage (*Dictionary of Computing*, page 110, 2nd paragraph). A complimentary excerpt of the cited reference has been provided to the Examiner as an Appendix I to this Amendment.

An index, on the other hand, is defined as a set of links that can be used to locate records in a data file. In a single-level index, an entry points directly to an individual record or to a group of records. At the higher levels of a multilevel index an entry points to a group of entries at the next lower level (*Dictionary of Computing*, page 217, next-to-last paragraph - Appendix I).

It is significant that an index, except for values of the particular data entity being indexed, does not in itself contain data from the database. It is also noteworthy that the internal layout of an index is not described by a schema in such a manner in the art that the index can be accessed or browsed as a standalone database; an index is not accessed using the facilities of a DBMS except as an index to associated data in a database. Without the associated data in the database, an index would be of little value.

An example of defining a table for a DB2 database is provided in *DB2, Concepts, Design, and Programming*, Prentice-Hall, 1989 of which a complimentary excerpt has been provided to the Examiner as an Appendix II to this Response. The example under Defining Tables (page 204) shows the definition of a database table named SUPPLIERS

containing database entities named SUPSUPP, NAME, ADDRESS and CODE, having the data types of SMALLINT, CHAR, VARCHAR and SMALLINT respectively. The listed entities thus form table rows described by the associated data types and respective lengths. This represents a data table that can be accessed in a variety of ways, with or without an index.

An example of defining an index, however, is distinctly different than defining a table. Three examples are provided in *DB2, Concepts, Design, and Programming* for creating indexes named XQUOT, XQUOT2 and XINV1 respectively (page 206 - Appendix II). It is readily apparent that the administrator or programmer only has the option of listing the table name and respective column names that are to be indexed. The administrator does not know the internal structure of the index; that is handled internally by DB2. The named database table may then be accessed via the index, however, there is no DBMS mechanism for accessing the index itself because the index does not hold database entities.

The present application distinguishes over prior art indexes by using a multi-dimensional database as an index. For example, the index is described as being sparse, with only N cells containing data (page 16, lines 5-7). The multi- dimensional database serving as an index is built and loaded with values (page 17, lines 1-3), and because the index is itself a database, it may be accessed and navigated independently by utilizing standard APIs as known in the art. Page 20, lines 23-31 describes a process of navigating the index, and connecting to an index, which defines patentably over prior-art methods which, to the contrary, navigate a connected database via links in an index, rather than navigating the index itself. The present application, therefore, describes an index that is used in a novel and efficient manner.

Applicants respectfully submit that the index taught in the Agrawal[1] patent is not a multi-dimensional database. Rather, the index mechanism taught in the Agrawal[1] patent is simply a point access method (PAM), such as an R-tree index (col. 4, lines 27-30). Although the PAM index of the Agrawal[1] patent may index data having any number of dimensions, the index itself is not a multi-dimensional database as shown by the *Dictionary of Computing* and *DB2, Concepts, Design, and Programming* references provided to the Examiner.

For at least the above reasons, it is respectfully submitted that dependent claims 56, 62 and 68 are patentably distinct over the references of record.

In rejecting claims 56, 62 and 68, the Examiner likened a step in U.S. Patent No. 6,094,651 to Agrawal, et al. (Agrawal[2]) for locating anomalies in a k-dimensional database to the index storing deviation values as described in the present application. Agrawal[2] teaches locating data anomalies in a k-dimensional database, including the step of associating a surprise value with each cell in a data cube and indicating a data anomaly when the surprise value associated with a cell exceeds a threshold value(col. 2, lines 38-43. Applicants respectfully submit that the locating of anomalies as taught by Agrawal[2] bears no resemblance to the deviations described in the present application (page 14, line 10 - page 17, line 6).

For at least the above reasons, it is respectfully submitted that dependent claims 56, 62 and 68 are patentably distinct over the references of record.

It is also respectfully submitted that new claims 73-87 are patentably distinct and unobvious over the references of record.

Conclusion

In view of the above amendments, comments, and arguments presented, applicants respectfully submit that all pending claims are patentably distinct and unobvious over the references of record.

Allowance of all claims and early notice to that effect is respectfully requested.

Respectfully submitted,

FAY, SHARPE, FAGAN,
MINNICH & McKEE, LLP

Michael E. Hudzinska
Reg. No. 34,185
1100 Superior Avenue,
Seventh Floor
Cleveland, OH 44114-2518
(216) 861-5582 (tel)
(216) 241-1666 (fax)

N:\IBMZ\200022\GTY0000184V001.doc